

HPC-Training for Beginners

Focus on users with Windows background

Johann Spies

University of Stellenbosch

2-3 March 2010



UNIVERSITEIT-STELLENBOSCH-UNIVERSITY
jou kennisvenoot-jour knowledge partner

Outline

From Windows to Linux



Outline

From Windows to Linux

Basic Linux Command line



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments

Using Gridengine



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments

Using Gridengine

Submitting a job



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments

Using Gridengine

Submitting a job

Parallel environments



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments

Using Gridengine

Submitting a job

Parallel environments

Samba



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments

Using Gridengine

Submitting a job

Parallel environments

Samba

Do's and Don'ts



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments

Using Gridengine

Submitting a job

Parallel environments

Samba

Do's and Don'ts

X-forwarding and qmon



Outline

From Windows to Linux

Basic Linux Command line

Rhashasta: where is what?

How to get help

Compilers and related environments

Using Gridengine

Submitting a job

Parallel environments

Samba

Do's and Don'ts

X-forwarding and qmon

Sun Secure Global Desktop



Log in at head001.sun.ac.za

Dear Koos,

I have registered you as user on head001.sun.ac.za. You can now ssh to `koos@head001.sun.ac.za` using your name and surname as password (including spaces).

...



Log in at head001.sun.ac.za

Dear Koos,

I have registered you as user on head001.sun.ac.za. You can now **ssh** to `koos@head001.sun.ac.za` using your name and surname as password (including spaces).

...



head001?



A cluster consisting of head001 and 21 nodes (node001 - node021). Each node have 16G RAM and two quadcore processors giving 8 cores per node.

head001: Submit host

nodes: Execution hosts



ssh?

What is **ssh**?
Secure shell.



ssh?

What is a **shell**?

The command interpreter used to pass commands to an operating system; so called because it is the part of the operating system that interfaces with the outside world. (Definition according to “The Free On-line Dictionary of Computing (30 January 2010) [foldoc]”)

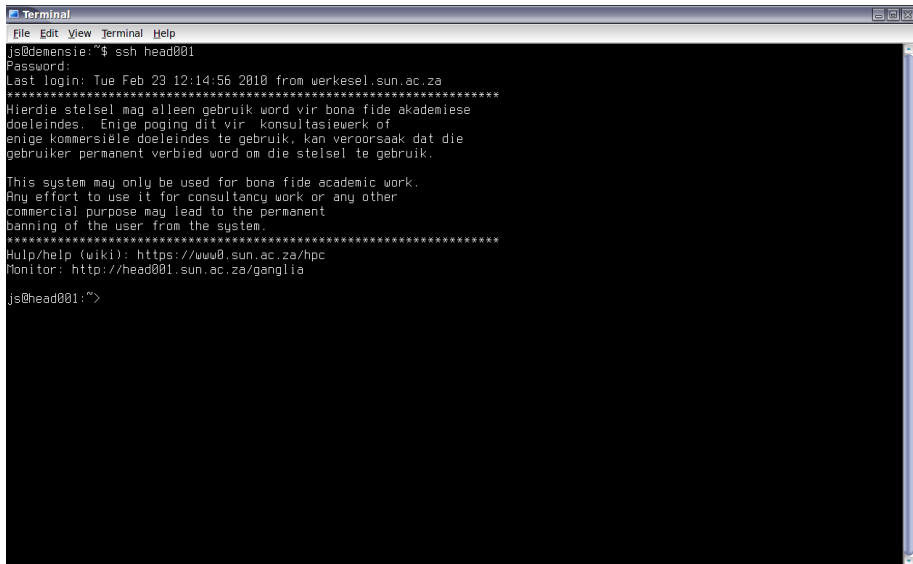


ssh?

An **ssh**-connection enables a user to communicate with the shell on the server using encrypted network traffic.



SSH



```
Terminal
File Edit View Terminal Help
js@demensie:~$ ssh head001
Password:
Last login: Tue Feb 23 12:14:56 2010 from werkesel.sun.ac.za
*****
Hierdie stelsel mag alleen gebruik word vir bona fide akademiese
doeleindes. Enige poging dit vir konsultasiewerk of
enige kommersiële doeleindes te gebruik, kan veroorsaak dat die
gebruiker permanent verbied word om die stelsel te gebruik.

This system may only be used for bona fide academic work.
Any effort to use it for consultancy work or any other
commercial purpose may lead to the permanent
banning of the user from the system.
*****
Hulp/help (wiki): https://www0.sun.ac.za/hpc
Monitor: http://head001.sun.ac.za/ganglia

js@head001:~>
```

SCP

A companion of ssh is [scp](#)(secure copy) which also known as [sftp](#):
A version of File Transfer Protocol (FTP) using an encrypted network connection provided by Secure Shell (SSH), usually SSH 2. (“The Free On-line Dictionary of Computing (30 January 2010) [foldoc]”)



Windows and Apple tools using ssh

- ▶ putty: Use it to log in on the remote system and get a shell prompt. Putty is not necessary when you use Linux or OS X.



Windows and Apple tools using ssh

- ▶ putty: Use it to log in on the remote system and get a shell prompt. Putty is not necessary when you use Linux or OS X.
- ▶ scp: (part of putty) from the command line to copy files over a secure network connection.



Windows and Apple tools using ssh

- ▶ putty: Use it to log in on the remote system and get a shell prompt. Putty is not necessary when you use Linux or OS X.
- ▶ scp: (part of putty) from the command line to copy files over a secure network connection.
- ▶ winscp: Use it to copy files between your computer and the server using the ssh connection.



Windows and Apple tools using ssh

- ▶ putty: Use it to log in on the remote system and get a shell prompt. Putty is not necessary when you use Linux or OS X.
- ▶ scp: (part of putty) from the command line to copy files over a secure network connection.
- ▶ winscp: Use it to copy files between your computer and the server using the ssh connection.
- ▶ Filezilla is an alternative to winscp and can be used on Linux, OS X and Windows.



Windows and Apple tools using ssh

- ▶ putty: Use it to log in on the remote system and get a shell prompt. Putty is not necessary when you use Linux or OS X.
- ▶ scp: (part of putty) from the command line to copy files over a secure network connection.
- ▶ winscp: Use it to copy files between your computer and the server using the ssh connection.
- ▶ Filezilla is an alternative to winscp and can be used on Linux, OS X and Windows.
- ▶ On OS X there is also Cyberduck that can be used for sftp.



Install putty and wincsp

► <http://ftp.sun.ac.za>



Install putty and wincsp

- ▶ <http://ftp.sun.ac.za>
- ▶ Search for putty.exe and install the latest version.



Install putty and wincsp

- ▶ <http://ftp.sun.ac.za>
- ▶ Search for putty.exe and install the latest version.
- ▶ Do the same for wincsp



Install putty and wincsp

- ▶ <http://ftp.sun.ac.za>
- ▶ Search for putty.exe and install the latest version.
- ▶ Do the same for wincsp
- ▶ Use putty to log in on head001.



Install putty and wincsp

- ▶ <http://ftp.sun.ac.za>
- ▶ Search for putty.exe and install the latest version.
- ▶ Do the same for wincsp
- ▶ Use putty to log in on head001.
- ▶ Do the same with wincsp



Commands often used

| Command | What does it do? | Windows/DOS |
|---------|--|-------------|
| ls | list directory | dir |
| cd | change directory | cd |
| cp | copy file to another place | copy |
| mv | move a file | ren or move |
| cat | print the contents of a text file | type |
| more | like cat with a paging function | more |
| less | like more but you can page up also | |
| find | find files | dir /s |
| df | How much free space on disk? | dir |
| mkdir | Make directory | mkdir or md |
| rmdir | Remove empty directory | rmdir or rd |
| grep | search for a string in a file | find |
| w | show who is logged in and what they do | |



Linux for Dos/Windows users

More about the equivalent commands/operators/programs can be found on the internet e.g.:

- ▶ http://www.yolinux.com/TUTORIALS/unix_for_dos_users.html
- ▶ <http://www.redhat.com/docs/manuals/linux/RHL-6.2-Manual/getting-started-guide/ch-doslinux.html>
- ▶ <http://www.faqs.org/docs/Linux-HOWTO/DOS-Win-to-Linux-HOWTO.html>



Do it

Try the following in your home directory:

- ▶ ls
- ▶ ls -la
- ▶ ls -lah
- ▶ ls -laR
- ▶ ls /
- ▶ ls /root
- ▶ ls -lapt
- ▶ ls -laprt (How does it differ from 'ls -lapt'?)
- ▶ man ls



Too much new information!

The following quote¹ from Paul Sheer's book, Rutebook, is important:

Any system reference will require you to read it at least three times before you get a reasonable picture of what to do. If you need to read it more than three times, then there is probably some other information that you really should be reading first. If you are reading a document only once, then you are being too impatient with yourself.

It is important to identify the exact terms that you fail to understand in a document. Always try to backtrack to the precise word before you continue.

Its also probably not a good idea to learn new things according to deadlines. Your UNIX knowledge should evolve by grace and fascination, rather than pressure.

¹my emphasis

Some useful shortcuts

TAB-key: Type `ls[TAB][TAB]`

Ctrl-r: Use Control-R to search your command history.

Cut & Paste: doubleclick on a word (or three times on a line) and use the middle (sometimes right) button to paste at the cursor.

Ctrl-t or Alt-t: type the following words in a line without pressing [Enter]:
one tow

Move the cursor (using the left and right arrows) to the 'w'
of 'tow'

Press Ctrl-t

Move the cursor to the 't' of 'two'

Press Alt-t

`cd [Enter]`] Takes you back to your home directory



Where to find help?

- ▶ What is the command?

Type 'apropos' and what you want to do e.g. `apropos compile | less`
or `apropos compile | grep fortran`

Then you can identify the command you want to use and use `man gfortran` to get the manual of that command.

- ▶ How do I use the command?

Use `man <command>` or `<command> --help`

- ▶ Documentation

Type `w3m /usr/share/doc` and explore the system.

- ▶ Questions

Send an email to `hpc@sympa.sun.ac.za` with your question. There is also another list (`sulug@sympa.sun.ac.za`) for the Stellenbosch Linux Users.



Try it...

date --help

Use **shift-pageup** or **shift-pagedown** to see the output or else use

date --help | less or date --help | more

or

date --help | less or date --help | less

less is more than more :)



Where is what on the system?

- ▶ Type 'env | less' and press Enter



Where is what on the system?

- ▶ Type 'env | less' and press Enter
- ▶ Type 'echo \$SGE_ROOT'



Where is what on the system?

- ▶ Type 'env | less' and press Enter
- ▶ Type 'echo \$SGE_ROOT'
- ▶ Type w3m and press Alt-. (This might not work in all cases e.g. when you ssh from a Macbook).



Where is what?

| | |
|-------------------|---|
| / | Also called 'root' |
| /export/home | The home directories of the users Mounted at the nodes. |
| /var | Where the system keeps files that change often |
| /bin, /usr/bin | Executables available to users |
| /sbin, /usr/sbin | Executables for the super user |
| /usr/local, /apps | Stuff that was not installed through the normal package system. Mounted at the nodes. |
| /opt | Some more stuff that was not installed using the package system. |
| /usr/share/doc | Documentation that comes with the standard packages |



Where is what?

Take some time, use `w3m` and “walk” around on the system. Ask questions.



How to get help

- ▶ <https://www0.sun.ac.za/hpc>
- ▶ <http://gridengine.sunsource.net> or the local copy:
<http://archive.sun.ac.za/ftp/pub/mirrors/gridengine.sunsource.net> for which you do not need Inetkey.
- ▶ <http://wikis.sun.com/display/gridengine62u3/Index>
- ▶ Send email to hpc@sympa.sun.ac.za
- ▶ Subscribe to the gridengine mailing list:
<http://gridengine.sunsource.net/maillist.html>
- ▶ Look at the examples in the 'examples'
- ▶ Read the man pages e.g. *man qsub*
- ▶ Read the manuals of the specific programs you are using.



Compilers and related environments

- ▶ **gcc:** GNU project C and C++ compiler



Compilers and related environments

- ▶ **gcc**: GNU project C and C++ compiler
- ▶ **sunstudio**: Sun Studio 12: C 5.9 compiler



Compilers and related environments

- ▶ **gcc**: GNU project C and C++ compiler
- ▶ **sunstudio**: Sun Studio 12: C 5.9 compiler
- ▶ **gfortran**: GNU Fortran compiler



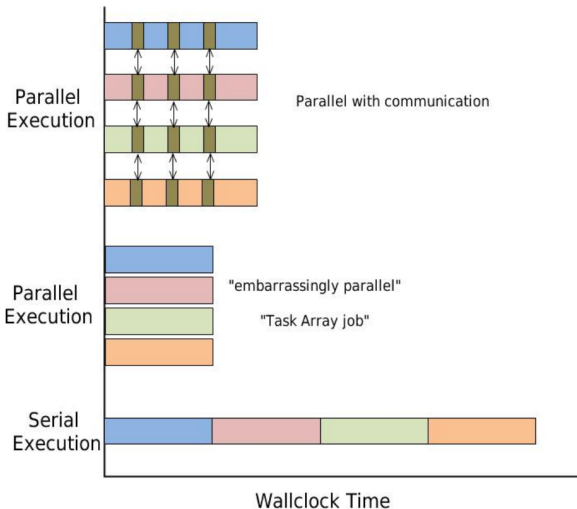
Compilers and related environments

- ▶ **gcc**: GNU project C and C++ compiler
- ▶ **sunstudio**: Sun Studio 12: C 5.9 compiler
- ▶ **gfortran**: GNU Fortran compiler
- ▶ Do the following:
 - ▶ `env > ~/xenv`
 - ▶ `which cc`
 - ▶ `module add sunstudio`
 - ▶ `env > ~/yenv`
 - ▶ `which cc`
 - ▶ `module avail`
 - ▶ `module rm sunstudio`
 - ▶ `sdiff ~/xenv ~/yenv | less`



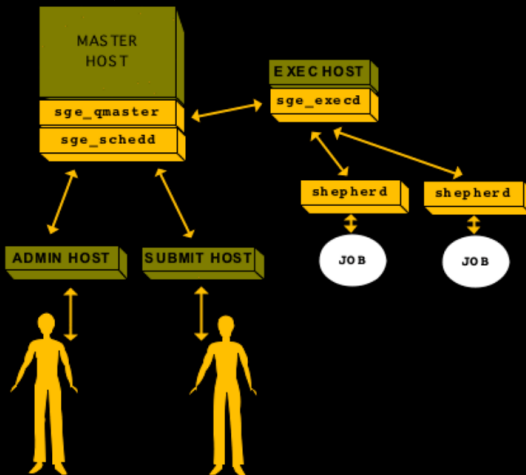
Terminology: Serial, Parallel

- Usually in HPC we run some kind of parallel workload
- Throughput jobs (independent tasks)
- Parallel programs



Host Types & SGE Daemons

- Exec hosts run 'sge_execd'
- Master runs 'sge_qmaster' 'sge_schedd' **
- Master optionally can run 'sge_execd'



```
#!/bin/sh
# Example job-script-file
# Set the name of the job.
#$ -N test
# Set the name of the stdout/stderr file
#$ -o serial.out
# Set the initial directory to be the current working directory
#$ -cwd
#Merge the standard out and standard error to one file
#$ -j y
#Inherit the current environment variables
#$ -V
# The max walltime for this job is 31 minutes
#$ -l h_rt=00:31:00
#Memory limit (1.3GB)
#$ -l h_vmem=1300M
./serial.exe
```



Do it!

► `cd ~/examples/1_sge_intro`



Do it!

- ▶ `cd ~/examples/1_sge_intro`
- ▶ Read the README files and submit `first.job`, `second.job` and `taskarray.job`



Do it!

- ▶ `cd ~/examples/1_sge_intro`
- ▶ Read the README files and submit `first.job`, `second.job` and `taskarray.job`
- ▶ `cd ../jobs`



Do it!

- ▶ `cd ~/examples/1_sge_intro`
- ▶ Read the README files and submit `first.job`, `second.job` and `taskarray.job`
- ▶ `cd ../jobs`
- ▶ change 'simple.sh' to save all output in the present directory in the file with the name `simple.out`. Submit `simple.sh` and check the output.



Time for problem solving...

- ▶ `cd ~/examples/3_parallel/simple_MPI`
- ▶ Read the README files and try compiling `primes_mpi.c` and submit the job. Try using a combination of `locate`, `grep`.
- ▶ How do you solve the problems you experience?
- ▶ Do `tail -f <outputfile>`



Time for problem solving...

- ▶ `cd ~/examples/3_parallel/simple_MPI`
- ▶ Read the README files and try compiling `primes_mpi.c` and submit the job. Try using a combination of `locate`, `grep`.
- ▶ How do you solve the problems you experience?
- ▶ Do `tail -f <outputfile>`
- ▶ Notice the usage of `'man qconf'`, `'locate'`, `'grep'`, `'export'` and `'echo'` `'qconf -spl'` `'qconf -sp'` in the problem solving.



Time for problem solving. . .

- ▶ `cd ~/examples/3_parallel/simple_MPI`
- ▶ Read the README files and try compiling `primes_mpi.c` and submit the job. Try using a combination of `locate`, `grep`.
- ▶ How do you solve the problems you experience?
- ▶ Do `tail -f <outputfile>`
- ▶ Notice the usage of 'man qconf', 'locate', 'grep', 'export' and 'echo' 'qconf -spl' 'qconf -sp' in the problem solving.
- ▶ Try a different PE's (Parallel Environment). What difference does it make?



Time for problem solving. . .

- ▶ `cd ~/examples/3_parallel/simple_MPI`
- ▶ Read the README files and try compiling `primes_mpi.c` and submit the job. Try using a combination of `locate`, `grep`.
- ▶ How do you solve the problems you experience?
- ▶ Do `tail -f <outputfile>`
- ▶ Notice the usage of 'man qconf', 'locate', 'grep', 'export' and 'echo' 'qconf -spl' 'qconf -sp' in the problem solving.
- ▶ Try a different PE's (Parallel Environment). What difference does it make?
- ▶ Try what is suggested in README.interactive



Time for problem solving. . .

- ▶ `cd ~/examples/3_parallel/simple_MPI`
- ▶ Read the README files and try compiling `primes_mpi.c` and submit the job. Try using a combination of `locate`, `grep`.
- ▶ How do you solve the problems you experience?
- ▶ Do `tail -f <outputfile>`
- ▶ Notice the usage of 'man qconf', 'locate', 'grep', 'export' and 'echo' 'qconf -spl' 'qconf -sp' in the problem solving.
- ▶ Try a different PE's (Parallel Environment). What difference does it make?
- ▶ Try what is suggested in README.interactive
- ▶ Try a different `module` environment: 'module avail' and 'module add' | 'module rm'



SSH and Parallel Environments

- ▶ `cd ~/.ssh`
- ▶ `ls -la`
- ▶ If there are no key-files: `ssh-keygen -t dsa`
- ▶ `ssh-copy-id -i ~/.ssh/id_dsa.pub comp001`
- ▶ If you are using something like Feko with a licence on each node it might be necessary to update your `known_hosts`-file by doing the following:

```
for i in {1..9}; do ssh comp00$i ls; done
for i in {10..21}; do ssh comp0$i ls; done
```

and answer 'y' for prompts asking you whether the node must be added to `known_hosts`.



Samba available

A Samba-service is available if you want to mount you home directory (or a subdirectory) on your own computer.

Configuration for such usage is handled on a per-request basis.



Do's and Don'ts

- ▶ Read the manuals and other documentation.
- ▶ Try again.
- ▶ Don't run your job on head001. Always use qsub to submit a job.
- ▶ Be patient.
- ▶ Kill your job (qdel) if you see that you have submitted a job by mistake or that your job is not behaving as expected and using too much of the resources.
- ▶ Don't start your work two weeks before it has to be handed in.
- ▶ Don't use the maximum available resources.



Getting a connection with graphical capabilities

- ▶ Install a X11-server on your Windows e.g. xming.



Getting a connection with graphical capabilities

- ▶ Install a X11-server on your Windows e.g. xming.
- ▶ Select the X-option in Putty



Getting a connection with graphical capabilities

- ▶ Install a X11-server on your Windows e.g. xming.
- ▶ Select the X-option in Putty
- ▶ Alternatively install [Sun VirtualBox](#) and use that to install a Linux Virtual Machine and use it to `ssh -X` to head001.



Sun Secure Global Desktop (available on Youtube)